



به نام هستی بخش

و سلام بر مهدی که انتظارش را نه فقط دل عاشق،  
که ترنم هر باران بهاری و هر روزنهی امید میکشد...

تمرین ۶ (فصل های ۸ و ۹)

موعد تحویل:

درس پایگاه داده‌ها، بهار ۱۴۰۳

۱) ابتدا توضیح دهید clustered یا unclustered بودن شاخص به چه معناست و اهمیت آن چیست؟ (۶ نمره) سپس در هر یک از سناریوهای زیر، توضیح دهید بهتر است از شاخص clustered استفاده شود یا unclustered یا فرقی ندارد؟ (۴ نمره)

- یک شاخص درختی alt. ۲ داریم و بیشترین پرس‌وجو از نوع equality search است.
- یک شاخص بر روی سن افراد می‌خواهیم تعریف کنیم و بیشترین پرس‌وجو از نوع range search بر روی سن افراد است. قبلاً یک شاخص clustered بر روی محل تولد تعریف کرده‌ایم.

۲) ۱- شاخص hash را توضیح دهید. بررسی کنید این شاخص برای equality search بهتر عمل می‌کند یا range search؟ (۵ نمره)

۲- تفاوت شاخص alt. ۲ و alt. ۳ را با رسم شکل توضیح دهید. مختصر توضیح دهید در هر پایگاه داده چند شاخص از نوع alt. 1, alt. ۲ و alt. ۳ می‌توانیم داشته باشیم؟ (۵ نمره)

۳- آیا استفاده از شاخص‌ها می‌تواند موجب کاهش سرعت عملیات بروزرسانی در پایگاه داده شود؟ با ذکر مثال توضیح دهید. (۵ نمره)

۳) ۱- دو روش page directory و linked list را برای پیاده سازی heap file در نظر بگیرید. تفاوت این دو روش را به اختصار توضیح دهید. (۵ نمره)

۲- در پایگاه داده‌ای ۱۸ صفحه داده داریم و ظرفیت بافر مورد استفاده نامحدود است. schema پایگاه داده بصورت زیر است و heap file استفاده می‌کنیم. برای پرسش داده شده، بررسی کنید در بدترین حالت به چند عملیات IO نیاز است؟ (۱۰ نمره)

Student(sId INTEGER, age INTEGER, name VARCHAR(20), email VARCHAR(20))

- Insert into Students(810100789, "name", "name@ut.ac.ir")

الف) پیاده سازی با page directory و هر دایرکتوری ظرفیت ۱۰ اشاره گر را دارد.

ب) استفاده از linked list و دوازده صفحه پر به همراه شش صفحه دارای فضای خالی.

۴) ۱- index-only-plan چیست؟ مزیت این روش را بیان کنید. (۲ نمره)

۲- با در نظر گرفتن رابطه زیر، برای هر یک از پرس‌وجوهای داده شده، بهترین کلید شاخص برای پاسخ index-only را مشخص کنید. (۶ نمره)

product(pid, name, price, category, rate)

a) select P.category, MIN(P.price) from product P  
where P.rate > 5  
group by P.category

b) select AVG(P.price) from product P  
group by P.category

۳- توضیح دهید در چه شرایطی می‌توان یک index-only-plan ارائه کرد؟ یک مثال ارائه دهید که این کار ممکن نباشد. (۴ نمره)

۵) روش‌هایی برای ذخیره رکوردها با طول متغیر بیان کرده و مختصراً توضیح دهید (بیان دو روش کافی است). یکی از این روش‌ها استفاده از offset است، مزایا و معایب این روش در برابر سایر روش‌ها را بیان کنید. (۷ نمره)

۶) پنج صفحه فیم در اختیار داریم. می‌خواهیم به ترتیب زیر و از چپ به راست به صفحات دستیابی داشته باشیم. استراتژی‌های LRU و MRU را در نظر بگیرید و برای هر یک، ترتیب اختصاص فریم‌ها به صفحات را مشخص کرده و در انتها hit rate را محاسبه کنید. تصور شود صفحات پس از دسترسی unpin شوند. (۱۰ نمره)

۵ - ۱ - ۴ - ۲ - ۳ - ۱ - ۲ - ۱ - ۶ - ۱ - ۲ - ۳ - ۵ - ۴ - ۶ - ۵ - ۲ - ۳ - ۱

۷) با در نظر گرفتن مفروضات زیر، هزینه range search و equality search را با فرض داشتن شاخص clustered hash index و clustered tree index و unclustered tree index محاسبه کنید. توجه شود محاسبات با فرض داشتن هر یک از این شاخص‌ها باید انجام شود و در نهایت برای شش حالت باید بررسی کرده باشید. (۱۰ نمره)

B: number of data pages.

R: number of records per page.

D: each IO time.

Alternative 2: data entry size = 10% of the record size.

40% capacity of pages are used.

۸) ۱- برای سطوح RAID بیان شده در درس، بیان کنید هر سطح چه ایرادی از سطح (یا سطوح) قبل را برطرف می‌کند؟ (۶ نمره)  
۲- چهار دیسک با ظرفیت نامحدود در اختیار دارید که متشکل از بلاک‌های ۸-بیتی می‌باشند. متن زیر را در این دیسک‌ها و با رعایت سطح RAID ۵ ذخیره کنید. توجه شود که  $P(A, B, C)$  بیانگر parity مربوط به A, B, C می‌باشد. همچنین توجه شود هر کاراکتر در extended ascii از ۸-بیت تشکیل شده است. (۵ نمره)

HELLO\_FROM\_THE\_OTHER\_SIDE

سربلند و پیروز باشید.